



## Detection and Isolation of Pectoral Muscle from Digital Mammogram: An Automated Approach

Indra Kanta Maitra\*

Research Fellow

Dept. of Computer Science & Engineering,  
University of Calcutta,  
Kolkata, India  
[ikm.1975@ieee.org](mailto:ikm.1975@ieee.org)

Sanjay Nag,

Research Fellow

Dept. of Computer Science & Engineering,  
University of Calcutta,  
Kolkata, India  
[sanjaynag75@gmail.com](mailto:sanjaynag75@gmail.com)

Samir K. Bandyopadhyay

Professor

Dept. of Computer Science & Engineering,  
University of Calcutta,  
Kolkata, India  
[skb1@vsnl.com](mailto:skb1@vsnl.com)

**Abstract:** Computer Aided Diagnosis (CAD) systems have improved diagnosis of abnormalities in mammogram images. Extraction of the breast region allows the search for abnormalities to be limited to the region of the breast without undue influence from the non-breast region. After performing essential pre-processing steps the exact breast region as the region of interest (ROI), has to be segmented. In this paper we propose a fully automated detection, isolation and segmentation of Pectoral Muscle from the breast region in mammographic images. This composite method have been implemented and applied to MIAS mammographic database using ground truth images and quantitative metrics to evaluate its performance characteristics.

**Keywords:** Mammogram, Gaussian Smoothing; Binary Homogeneity Enhancement Algorithm (BHEA); Edge Detection Algorithm (EDA); Pectoral Muscle Detection Algorithm (PMDA); Pectoral Boundary Smoothing Algorithm (PBSA).

### I. INTRODUCTION

Breast cancer is the second largest cause of death of women all over the world [1]. Its causes are still not understood. So, the most effective approach to reducing the mortality from breast cancer is through early detection. Routine screening of women over the age of 40 with X-ray mammography has been demonstrated to contribute to mortality reduction. Current methods of treatment are very effective against breast cancer in its early phase, when the balance between the tumor and its host is more favorable [2]. Therefore, removal of cancer while it is still in early stages is the most promising way to achieve a significant change in the current breast cancer situation [3]. The large number of mammograms generated by population screening must be interpreted and diagnosed by relatively few radiologists. It is considered that the use of computerized mammographic analysis will make a vital contribution to easing the increasing workload and assisting in the detection of breast cancer.

The pectoral muscle a non-breast region in mammograms acts like an additional complexity in automated analysis using Computer Aided Diagnosis (CAD) systems. The pectoral muscle represents a predominant density region in most medio-lateral oblique (MLO) views [4] of mammograms, and can affect the results of image processing methods. Intensity-based methods can present poor results when applied to differentiate dense structures such as suspicious masses or fibro-glandular discs, because the pectoral muscle has a similar opacity as tumors. Another important reason for detecting the pectoral muscle in mammograms is that we can detect the possible presence of abnormal axillary lymph nodes, which is a major sign of breast carcinoma. This research paper presents a new and

efficient segmentation method using proposed binary homogeneity algorithm with color quantization followed by a noble edge detection algorithm and a unique line traversing algorithm to delineate pectoral muscle region accurately. The proposed method is capable of segmenting pectoral muscle of a broad range of size, shape, type and position. To demonstrate the capability of segmentation algorithms that is extensively tested on all the mammograms of MIAS database [5] using ground truth images and quantitative metrics to evaluate its performance characteristics.

### II. LITERATURE REVIEW

There have been various approaches to the task of segmenting the pectoral muscle in mammograms. Suckling et al. [6] segmented mammograms into four major components: background, pectoral muscle, fibro-glandular region and adipose region, using multiple, linked self-organizing neural networks. Aylward et al. [7] used a gradient magnitude ridge traversal algorithm at a small scale to extract multiple initial points and then resolved the resulting multiple edge definitions via a voting scheme. Saha et al. [8], [9] reported a semi-automatic method that requires input from an operator to locate the pectoral muscle; delineation is then performed automatically. Chandrasekhar and Attikiouzel proposed two techniques to enhance the pectoral muscle region on mammograms: the extended Russ operator [10], and tunable parametric edge detection [11], although final segmentation on the enhanced images had not been carried out.

The Hough transform, used by Karssemeijer [12] to detect the pectoral edge, is a popular technique. This method assumes that the pectoral edge is approximately a straight line oriented in a certain direction. To ensure that the correct peak is selected in the Hough space, gradient magnitude and

orientation, length of projected line, and corresponding pectoral area were taken into account [12]. Other recent studies based on the Hough transform include: Ferrari et al. [13] who segmented mammograms into skin-air boundary, fibro-glandular tissue, and pectoral muscle; Yam et al. [14] who refined the Hough transform linear approximation into a curved pectoral boundary using a dynamic programming method; and Georgsson [15] who extracted the pectoral muscle by region growing.

In summary, almost all of the previous work approximates the pectoral edge as a straight line and then refines that straight line into a more accurate curved boundary. The accuracy of these algorithms is usually good. With the Hough transform, the pectoral muscle was segmented in 92.8% of the mammograms tested in Karssemeijer; and the pectoral muscle was accurately detected in 87.9% of the images within a difference of 1–3 mm in Ferrari et al.

### III. PROPOSED METHODS

Digital Mammograms are medical images requires a preparation phase in order to improve the image. Our objective during this process is to improve the quality of the image to make it ready for further processing by removing the irrelevant and unwanted parts in the background of the mammogram.

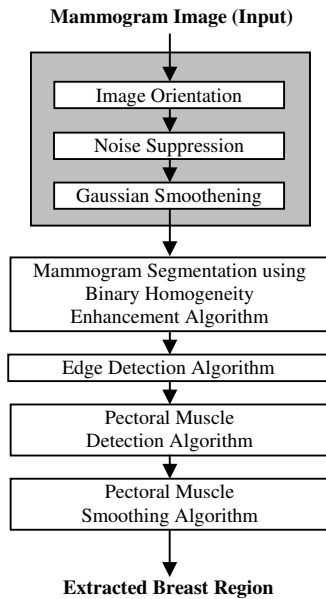


Figure 1. Proposed Algorithm Flowchart

#### A. Mammogram Preprocessing

##### 1) Image Orientation

The mammogram image is transformed, so that the chest wall location, i.e., the side of the image containing the pectoral muscle, is on the upper left corner of the image. In order to determine the chest wall location, the decreasing pixel intensity of the breast tissue near the skin-air interface is used. This tissue is estimated by employing the Minimum cross-entropy threshold selection technique [16] twice in the original image. After determining the chest wall location we determine the top of the image. We extract the vertical centroid of the image and then we assume that the asymmetric region closest to the right side of the vertical centroid is the tip of the breast. The image is flipped

horizontally if needed to place the asymmetric region below the vertical centroid, resulting in an image that is right way up.

##### 2) Noise Suppression

There are different types of noises, which appear in MIAS images. High intensity noise is characterized by high values of optical densities, such as labels or scanning artifacts. Tape artifacts are markings left by tapes, or other shadows presenting themselves as horizontal running strips. Such noise must be replaced by black pixels.

##### 3) Gaussian Smoothing

The Gaussian smoothing operator is a 2-D convolution operator that is used to 'blur' images and remove detail and noise. It uses a kernel that represents the shape of a Gaussian ('bell-shaped') hump. In 2-D, an isotropic (i.e. circularly symmetric) Gaussian has the form:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

The idea of Gaussian smoothing is to use this 2-D distribution as a 'point-spread' function, and this is achieved by convolution. Since the image is stored as a collection of discrete pixels we need to produce a discrete approximation to the Gaussian function before we can perform the convolution. In theory, the Gaussian distribution is non-zero everywhere, which would require an infinitely large convolution kernel, but in practice it is effectively zero more than about three standard deviations from the mean, and so we can truncate the kernel at this point. Once we obtain a suitable kernel, then the Gaussian smoothing can be performed using standard convolution methods.

In our proposed method we have taken 7×7 kernel as a convolution filter. The mammogram images in the MIAS database are categorized broadly into 3 categories namely Fatty, Fatty-Glandular and Dense-Glandular depending on the density of fatty tissues and abundance of glands in the breast. Close observation of all the mammograms reveal that each category of mammogram displayed a varied intensity value which is distinct from each category. This property of the mammograms has helped us in our choice of the value of deviation ( $\Omega$ ) for each category, thus able to adjust the level of smoothing for each category.

- Step 1. Calculate Gaussian Kernel values to create a convolution filter.
- Step 2. Convolution of the mammogram image starting from Left Topmost pixel to Bottom Rightmost pixel.
- Step 3. Perform matrix multiplication between calculated Gaussian filter values and pixel intensity of convolved image.

#### B. Pectoral Muscle Detection and Isolation

This paper is based on the image segmentation method, the inputs are images and, outputs are the attributes extracted from those images. Segmentation divides image into its constituent regions or objects. Segmentation plays an important role in image analysis. The goal of segmentation is to isolate the regions of interest (ROI) depending on the problem and its characters. The approach is to partition an image based on abrupt changes in intensity, such as edges in an image and partitioning image into regions that are similar according to a set of predefined criteria.

A mammogram contains two distinctive regions, the exposed breast region and the unexposed air-background (non-breast) region. The principal feature on a mammogram

is the breast contour, otherwise known as the skin-air interface, or breast boundary and the pectoral muscle. The breast region can be obtained by partitioning the mammogram into breast, non –breast region and pectoral muscle. On isolation of both non-breast and pectoral muscle region we can obtain the breast region for further studies. This paper proposes a new algorithm to identify and isolate pectoral muscle from the breast region.

#### 1) Binary Homogeneity Enhancement Algorithm (BHEA)

In this paper we have proposed a new homogeneity enhancement process namely Binary Homogeneity Enhancement Algorithm (BHEA) for digital mammogram. According to this method we take a maximum difference threshold (MDT) value, which is constant threshold determine by observation. We start checking this value with the image data by horizontally scanning from left of the array to the right from the topmost row of the mammogram image. If result of any subtraction is greater than the MDT, the array will be divided into two equal subsets along middle position and the first and last positions of the two subsets will be pushed to stack. Otherwise, the mode value of subset will be propagated to all other position after modifying value using uniform color quantization technique in color space breaking it into sixteen level scales. The process will be continued recursively, checking the stack, till the rightmost pixel in the last row of the mammogram image is reached. The same process will be repeated by scanning the image vertically from top to bottom followed by uniform color quantization. Algorithm for Horizontal Processing:

- Step 1. Scan Image from first row from left most pixel to right most pixel.
- Step 2. Check if the value of pixel is:  $MDT-val < pixel\ intensity < MDT+val$ .
- Step 3. If value differs then divide the array from start to end through the midpoint. Push the Start and End Position of both halves into the Stack and Goto Step2.
- Step 4. If values of consecutive pixels satisfy the condition then continue till the value differs. Then find Mode value of all the pixels of similar intensity and replace them in the image after performing uniform color quantization for the set of pixels.
- Step 5. Check Stack. If not empty, Pop value from Stack and Goto Step2.
- Step 6. Continue till rightmost pixel of the Last row is reached.

#### Algorithm for Vertical Processing:

- Step 1. Scan Image from first column from top most pixels to bottom most pixels.
- Step 2. Check if the value of pixel is:  $MDT-val < pixel\ intensity < MDT+val$ .
- Step 3. If value differs then divide the array from start to end through the midpoint. Push the Start and End Position of both halves into the Stack and Goto Step2.
- Step 4. If values of consecutive pixels satisfy the condition then continue till the value differs. Then find Mode value of all the pixels of similar intensity and replace them in the image after performing uniform color quantization for the set of pixels.
- Step 5. Check Stack. If not empty, Pop value from Stack and Goto Step2.
- Step 6. Continue till bottom most pixel of the Last column is reached.

#### 2) Edge Detection Algorithm (EDA)

In this research paper we have proposed a new method for detecting edge rather than using one of the known methods of edge detection. It can be observed that all the mammogram images in the mini MIAS database, have images of the breast at the center (left to right or right to left), and is bounded by non-breast image on either side. The mammogram image have skin-air interface on the side of the nipple where the air interface have black pixels. Similarly, the other side behind the breast has a well-defined baseline that separates the breast image from the body and contains black pixels. We consider these features in our proposed method of edge detection. In the first step of our proposed method we perform horizontal scanning. If any change of pixel intensity is observed it is marked by a black pixel indicating a horizontal edge point. We continue this process for all rows of pixel data to obtain a Horizontal Edge Map. In the next step, we scan the image vertically. Continuing the process for all the columns we obtain a Vertical Edge Map image. Finally, we merge the Horizontal Edge Map with Vertical Edge Map by performing a logical OR operation on the two image files, to obtain the Edge map of mammogram image. Algorithm for Horizontal Image Map:

- Step 1. Scan the Image Array Horizontally from left-most pixel to right-most pixel from first row to last row.
- Step 2. Take the first pixel intensity value as a reference value.
- Step 3. Compare intensity of subsequent pixels with the reference value. If the value is same continue to next pixel.
- Step 4. If the value differs, change the value of reference value to the pixel intensity value and mark the pixel black.
- Step 5. If the last row and column pixel is not reached then Goto Step3.

#### Algorithm for Horizontal Image Map:

- Step 1. Scan the Image Array Vertically from top-most pixel to bottom-most pixel from first column to last column.
- Step 2. Take the first pixel intensity value as a reference value.
- Step 3. Compare intensity of subsequent pixels with the reference value. If the value is same continue to next pixel.
- Step 4. If the value differs, change the value of reference value to the pixel intensity value and mark the pixel black.
- Step 5. If the last column and row pixel is not reached then Goto Step3.

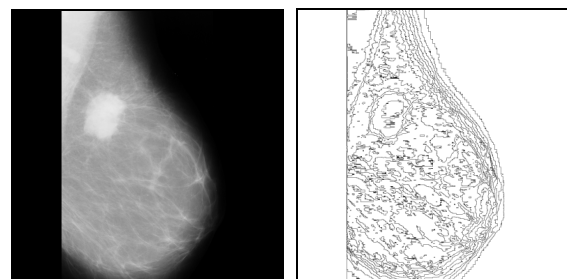


Figure 2. Raw Mammogram and Edge Map (MIAS 184.L)

#### 3) Pectoral Muscle Detection Algorithm (PMDA)

After we detect the edges of the mammogram by our proposed method, we get an image that contains the breast region clearly differentiated from the background. The breast region contains the outline edges of various breast constituents bounded by the breast boundary that defines the breast region from the image background. For further processing it is of utmost importance to extract the breast region by eliminating the pectoral muscle region. In this research paper we now proposed a new method for detection, extraction and isolation of Pectoral Muscle from the breast region.

A careful and close study of all the mammogram images in the MIAS database, it is observed that the pectoral muscle have a distinct intensity value that differs from the rest of the fatty tissues on the mammogram image. Similarly, it is also observed that the pectoral muscle lies at the top left corner of the breast region (Right pair flipped horizontally during preprocessing) and it traverses from the top margin downward towards the left baseline of the mammogram image forming an inverted triangle. Our objective is to identify this triangle from the breast region.

We now need to identify the outermost edge line that constitutes the edge of the Pectoral muscle. It is observed that the largest inverted triangle in the edge map starting from the top margin and ending on the vertical baseline, of the breast region on the left side, is the pectoral region. We output the pectoral boundary edge line on another image. We first identify the rightmost pixel of the breast region on the right side and draw a vertical line from top margin to bottom margin, parallel to the left vertical baseline. We draw another line parallel to the top and bottom margin passing through the rightmost pixel of the breast region on the right vertical line and ending on the left vertical baseline of the breast. We limit our search to the obtained rectangle thus reducing the processing time that would have incurred if we considered the entire mammogram. It is also observed that almost all, barring a few mammograms, have the pectoral region confined within this rectangle. Algorithm for Pectoral Muscle Isolation:

- Step 1. Scan the image from the right side of the image to locate the rightmost pixel of the breast region.
- Step 2. Draw a vertical line along this pixel from top to bottom.
- Step 3. Draw a horizontal line parallel to the top margin passing through the rightmost pixel on the right vertical line to the left base line.
- Step 4. Scan the edge map from the right side to left, on the obtained rectangle, from the first row.
- Step 5. Obtain a pixel that is black indicating an edge path, traverse the pixel path by considering all the surrounding pixels in a clockwise priority and consider the pixel with the highest priority.
- Step 6. The pixels that surrounded the edge pixel, but are of lower priority are stored in a Backtrack Stack to be used only if the traversal process reaches a dead end.
- Step 7. If a dead end is reached, pop out from the Backtrack stack a lesser priority pixel and continue with the traversal process.
- Step 8. Store the pixels traversed in a Plotting List to be used later for drawing the pectoral boundary.
- Step 9. Traversal continues to the next pixel till it reaches the left baseline or the bottom of the rectangle.
- Step 10. If the bottom of the rectangle is reached the path is discarded, the plot list is erased and continue from Step5.

- Step 11. Else the path indicated by the Plotting List is plotted on another image indicating pectoral edge.

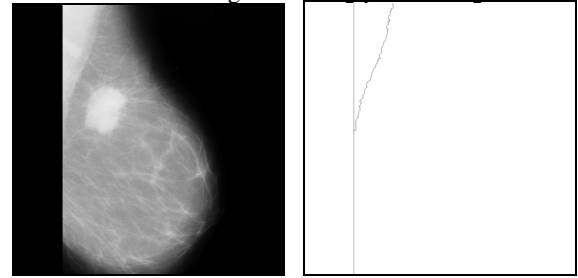


Figure 3. Raw Mammogram and Isolated Pectoral Boundary (MIAS 184.L)

#### 4) Pectoral Boundary Smoothing Algorithm (PBSA)

The Pectoral boundary image we obtain is further processed by our Pectoral Boundary Smoothing method. The pectoral boundary image obtained so far is not smooth. We obtain a smooth pectoral boundary by taking pixels at a fixed discrete interval and joining them by drawing a simple curve between the pixel positions to obtain a smooth and enhanced pectoral boundary. This forms our final output image. Algorithm for Pectoral Muscle Smoothing:

- Step 1. Get pixels from the Plotting List at a discrete interval.
- Step 2. Plot the first pixel.
- Step 3. Draw a simple curve between two consecutive pixel positions separated by the discrete interval.
- Step 4. Continue plotting till the last pixel in the Plotting List is plotted.

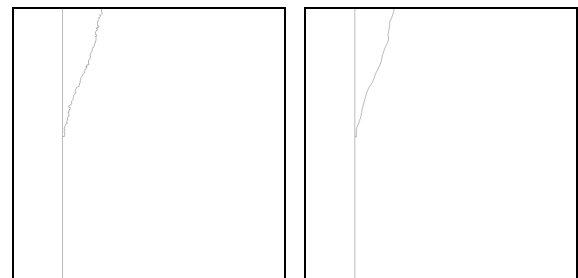


Figure 4. Detected Pectoral Boundary and Smooth Boundary (MIAS 184.L)

## IV. EXPERIMENTAL RESULTS EVALUATION

### A. Experimental Results

Results obtained after execution of our proposed algorithm on MIAS database Mammogram image mdb020.L.

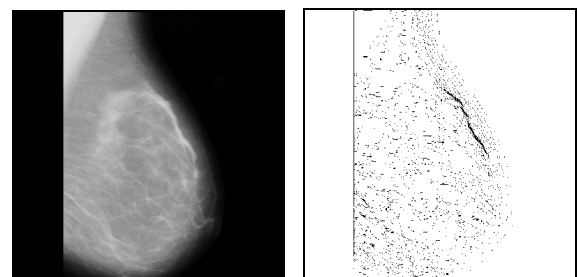


Figure 5. Raw Mammogram and Corresponding Edge Map

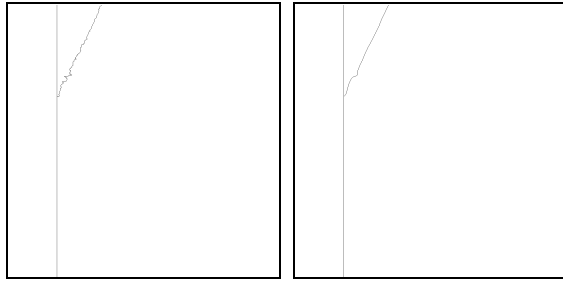


Figure 6. Detected Pectoral Boundary and Smooth Boundary

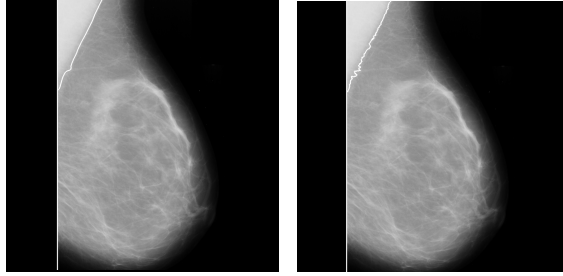


Figure 7. Detected Pectoral Boundary and GT Superimposed on Raw Mammogram

Results obtained after execution of our proposed algorithm on MIAS database Mammogram image mdb044.L.

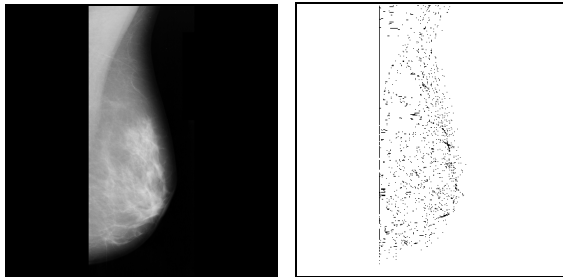


Figure 8. Raw Mammogram and Corresponding Edge Map

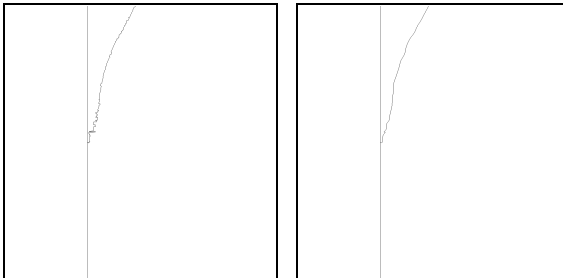


Figure 9. Detected Pectoral Boundary and Smooth Boundary

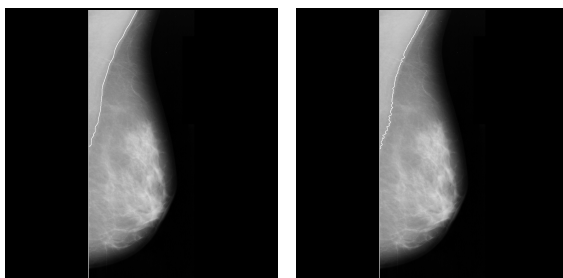


Figure 10. Detected Pectoral Boundary and GT Superimposed on Raw Mammogram

### B. Accuracy Estimation

Performance evaluation in algorithm design is an important step that is commonly neglected. Few mammogram segmentation algorithms have been tested extensively. What constitutes an “acceptable” result differs significantly, and is often based on visual subjective opinion with very little quantitative endorsement. The accuracy of this technique was evaluated through quantitative measures derived through the comparison of each segmented mammogram “mask” with its corresponding “gold standard”. The gold standard is generated by manually segmenting the breast region from each mammogram. The boundary of the breast is then manually traced to extract the real breast region to generate a ground truth (GT) image.

A quantitative measure was then derived to describe the accuracy of the segmentation. The region extracted by the segmentation algorithm (mask), which matches the GT, is denoted as true positive (TP) emphasizing that the algorithm has indeed found a portion of the breast. Pixels shown in the GT but not shown in the mask are defined as false negative (FN) classifications. These are considered missing pixels in the pectoral region. Conversely, the pixels not in the GT, but in the mask are defined as false positive (FP) pixels. By using these parameters on the output of all the 322 mammogram images in MIAS database, we obtain the following result on different quality measures.

Table I. Common measures used in the evaluation of our proposed methods

Common Measures	Computation	Mean Result
Accuracy (Percentage agreement)	$\frac{ TN + TP }{ TN + TP + FP + FN }$	0.9983
Dice similarity coefficient (DSC)	$\frac{2 \times  TP }{2 \times  TP + FP + FN }$	0.9508
Error rate	$\frac{ FP + FN }{ FP + FN + TP + TN }$	0.0016
Sensitivity (Percentage of Correct Estimation)	$\frac{ TP }{ TP + FN }$	0.9897
Specificity (True Negative Fraction/Rate)	$\frac{ TN }{ TN + FP }$	0.9602
False Positive Fraction/Rate	$1 - \text{Specificity}$	0.0397
Under estimation fraction (UEF)	$\frac{ FN }{ TN + FN }$	0.0003
Over estimation fraction (OEF)	$\frac{ FP }{ TN + FN }$	0.0013

### V. CONCLUSION

In our proposed method, we have projected four distinct steps. The ultimate smooth pectoral boundary image we obtain by the proposed Binary Homogeneity Enhancement Algorithm (BHEA), followed by Edge Detection Algorithm (EDA) and Pectoral Muscle Detection Algorithm (PMDA), and finally proposed Pectoral Boundary Smoothing Algorithm. Results of the proposed method show a reliable detection rate of pectoral muscle, furthermore, due to its simple procedure the method executes faster than other complicated methods. Performance evaluation in algorithm design is a commonly neglected concept. Few mammogram segmentation algorithms have been tested extensively. Our proposed method has been tested on all 322 mammograms from mini-MIAS mammogram database. This method can be used for other medical image processing applications with minor modification. In conclusion, it can be mentioned, our

proposed method is acceptably accurate, promising and comparable with any other standard methods.

## VI. REFERENCES

- [1] S. H. Landis, T. Murray, S. Bolden and P. A. Wingo, "Cancer Statistics", CA Cancer Journal, vol. 48, pp. 6-29, 1998
- [2] F. Winsberg, "Detection of radiographic abnormalities in mammograms by means of optical scanning and computer analysis", Radiology, vol. 89, pp. 211-215, 1967
- [3] S-M. Lai, X. Li, W.F. Bischof, "On techniques for Detecting Circumscribed Masses in Mammograms", IEEE Transactions on Medical Imaging, vol. 8, no 4, pp. 377-386, 1989
- [4] R. M. Rangayyan, Biomedical Image Analysis, IEEE Press/Wiley, New York, 2002
- [5] Suckling J., Parker J., Dance D.R., Astley S., Hutt I., Boggis C.R.M., Ricketts I., Stamatakis E., Cernaez N., Kok S.L., Taylor P., Betal D., Savage J., "The Mammographic Image Analysis Society Digital Mammogram Database", Proceedings of the 2nd International Workshop on Digital Mammography, York, England, 10-12 July 1994, Elsevier Science, Amsterdam, The Netherlands, pp. 375-378, 1994
- [6] J. Suckling, D. R. Dance, E. Moskovic, D. J. Lewis, and S. G. Blacker, "Segmentation of mammograms using multiple linked self-organizing neural networks," Med. Phys., vol. 22, no. 2, pp. 145-152, 1995
- [7] S. R. Aylward, B. M. Hemminger, and E. D. Pisano, "Mixture modeling for digital mammogram display and analysis," in Digital Mammography: Nijmegen, 1998, N. Karssemeijer, M. Thijssen, J. Hendriks, and L. van Erning, Eds. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1998, vol. 13, Computational Imaging and Vision, pp. 305-312, 1998
- [8] P. K. Saha, J. K. Udupa, E. F. Conant, D. P. Chakraborty, and D. Sullivan, "Breast tissue density quantification via digitized mammograms," IEEE Trans. Med. Imag., vol. 20, pp. 792-803, 2001
- [9] P. K. Sinha, J. K. Udupa, E. F. Conant, and D. P. Chakraborty, "Near automatic quantification of breast tissue glandularity via digitized mammograms," in Proc. SPIE: Medical Imaging 1993, K. M. Hanson, Ed. San Diego, CA, USA: SPIE, 1999, vol. 3661, SPIE Proceedings, pp. 266-276, 1999
- [10] R. Chandrasekhar and Y. Attikiouzel, "New range-based neighborhood operator for extracting edge and texture information from mammograms for subsequent image segmentation and analysis," IEE Proceedings—Science, Measurement and Technology, vol. 147, no. 6, pp. 408-413, 2000
- [11] "Segmentation of the pectoral muscle edge on mammograms by tunable parametric edge detection," in Advances in Signal Processing and Computer Technologies, G. Antoniou, N. Mastorakis, and O. Panfilov, Eds. Athens, Greece: World Scientific and Engineering Society WSES Press, pp. 55-60, 2001
- [12] N. Karssemeijer, "Automated classification of parenchymal patterns in mammograms," Phys. Med. Biol., vol. 43, no. 2, pp. 365-378, 1998
- [13] R. J. Ferrari, R. M. Rangayyan, J. E. L. Desautels, and A. F. Frère, "Segmentation of mammograms: Identification of the skin boundary, pectoral muscle, and fibroglandular disc," in IWDM 2000: Proc. 5th Int. Workshop Digital Mammography, M. J. Yaffe, Ed., Madison, WI, pp. 573-579, 2001
- [14] M. Yam, M. Brady, R. Highnam, C. Behrenbruch, R. English, and Y. Kita, "Three-dimensional reconstruction of microcalcification clusters from two mammographic views," IEEE Trans. Med. Imag., vol. 20, pp. 479-489, 2001
- [15] F. Georgsson, "Algorithms and techniques for computer aided mammographic screening," Ph.D. dissertation, Umeå Univ., Dept. Comput. Sci., Umeå, Sweden, 2001
- [16] A.D.Brink and N.E.Pendock, "Minimum Cross-entropy Threshold Selection, Pattern Recognition", 29(1): 179-188, 1996